Characterization of the core microbial community across different aggregate sizes in full-scale aerobic granular sludge plants and their relevance to wastewater treatment performance

Ruiz-Haddad, Lucia; Shaw, Dario Rangel; Ali, Muhammad; Pronk, Mario; van Loosdrecht, Mark C.M.; Saikaly, Pascal E.

DOI
10.1016/j.watres.2024.123036

Publication date
2025

Document Version
Final published version

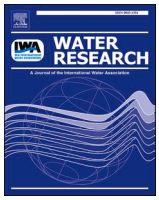
Published in
Water Research

Water Research 274 (2025) 123036

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres



Characterization of the core microbial community across different aggregate sizes in full-scale aerobic granular sludge plants and their relevance to wastewater treatment performance

Lucia Ruiz-Haddad [a,b], Dario Rangel Shaw [b], Muhammad Ali [c], Mario Pronk [d], Mark C.M. van Loosdrecht [d], Pascal E. Saikaly [a,b,*]

[a] Environmental Science and Engineering Program, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia
[b] Biological and Environmental Science and Engineering (BESE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia
[c] Department of Civil, Structural & Environmental Engineering, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland
[d] Department of Biotechnology, Delft University of Technology, Delft 2629 HZ, the Netherlands

ARTICLE INFO

Keywords:
Wastewater treatment
Aerobic granular sludge
Core community
Microbial aggregates
Nitrifying bacteria
Polyphosphate-accumulating organisms

ABSTRACT

Aerobic granular sludge (AGS) technology holds great promise of becoming the standard for biological wastewater treatment due to its lower energy consumption, small footprint, and high removal efficiency of nutrients compared to the conventional activated sludge processes. Different-sized aggregates have been shown to harbor a different microbial community composition. The central question is do full-scale AGS wastewater treatment plants (WWTPs) select for core microbial communities across different aggregate sizes and how these selected organisms differ between the different-sized aggregates. This study analyzed samples from nine geographically distributed full-scale AGS WWTPs that consistently perform well in terms of chemical oxygen demand (COD) and nutrient (N and P) removal. The main results showed that site-specific conditions highly influence microbial composition in smaller aggregates (< 1 mm), while larger granules form stable communities independent of WWTP location. Notably, all aggregates contained a small subset of 128–139 core OTUs that were both prevalent and abundant across all sizes. These core OTUs include key functional groups such as fermenters, aerobic heterotrophs, polyphosphate-accumulating organisms (PAOs), glycogen-accumulating organisms (GAOs), and nitrifiers, which play a crucial role in COD and nutrient removal. Additionally, an enrichment pattern was observed, with aerobic heterotrophs dominating in flocs, PAOs in small granules, and GAOs and nitrifiers in large granules. This study offers valuable insights into the core microbiome of different-sized aggregates in full-scale AGS WWTPs and highlights their potential role in overall system performance.

1. Introduction

Biological wastewater treatment plants (WWTPs) are the most widely implemented technology for treating wastewater (Sravan et al., 2024). They play a vital role in removing pollutants from wastewater, safeguarding the environment, and protecting public health (Ali et al., 2019). The most widely applied biological process in full-scale WWTPs is the activated sludge (AS) process, which relies on microbial aggregates (i.e., flocs) to effectively remove organic substrates and nutrients (N and P) from wastewater (Kim et al., 2019). Numerous studies have been conducted on full-scale AS technology to understand the composition and interactions of microbial communities in flocs to improve the efficiency and performance of WWTPs (Begmatov et al., 2022; Dueholm et al., 2022; Guo et al., 2017; Saunders et al., 2016; Wu et al., 2019). However, the aerobic granular sludge (AGS) process has emerged as a promising alternative to AS due to its superior nutrient removal capabilities and potential reductions of up to 80% in land footprint, 30% in energy consumption, and 50% in operational and capital costs (Pronk et al., 2015). These advantages have led to a rapid worldwide uptake of the technology in the past decade (Ali et al., 2023).

The AGS technology employs constant volume sequencing batch reactors with alternating anaerobic/anoxic and aerobic conditions, which promote the formation of stable microbial aggregates within the reactor (De Kreuk et al., 2005). These aggregates are generally classified

* Corresponding author.
E-mail address: pascal.saikaly@kaust.edu.sa (P.E. Saikaly).

https://doi.org/10.1016/j.watres.2024.123036
Received 24 October 2024; Received in revised form 15 December 2024; Accepted 22 December 2024
Available online 24 December 2024

as flocs (< 0.2 mm) and granules (> 0.2 mm), with granules comprising approximately 80% of the biomass (Ekholm et al., 2022; Pronk et al., 2015). The difference in aggregate sizes directly influences settling velocity during the settling phase, resulting in flocs accumulating in the upper layers and larger granules settling at the bottom of the reactor (Van Dijk et al., 2022). Consequently, flocs are more frequently discharged and exposed to lower substrate concentrations during the feeding period, which occurs from the bottom of the reactor (Ali et al., 2019). In contrast, granules benefit from longer solids retention time (SRT) and higher substrate concentrations (Ali et al., 2019). The difference in SRT, substrate concentration, and size-related microenvironments result in unique microbial compositions within each aggregate size fraction (Ali et al., 2019; Dueholm et al., 2022; Winkler et al., 2018). For instance, flocs in AGS are more susceptible than granules to harbor fast-growing microorganisms responsible for chemical oxygen demand (COD) removal (Ali et al., 2019; Aonofriesei and Petrosanu, 2007; Nielsen et al., 2004). Conversely, granules behave as biofilms supporting the coexistence of aerobic and anoxic species through the establishment of a concentration gradient of organic substrates, nutrients, and electron acceptors (e.g., O_2, nitrate, nitrite) (Winkler et al., 2018; Xia et al., 2018). Moreover, the establishment of concentration gradient, enhanced substrate concentration, and extended SRT support the proliferation of slow-growing organisms in granules (Dueholm et al., 2022; Ali et al., 2019; Winkler et al., 2018). These slow-growing organisms include functional groups of bacteria such as nitrifiers, polyphosphate-accumulating organisms (PAOs), and glycogen-accumulating organisms (GAOs) (Hamza et al., 2018; Winkler et al., 2013). Furthermore, granules have demonstrated greater resilience against the immigration of microbial communities from influent wastewater compared to flocs, suggesting enhanced microbial selection and stability (Ali et al., 2019; Dueholm et al., 2022; Winkler et al., 2018).

Previous microbial community studies of full-scale AGS WWTPs have mainly focused on: i) exploring the relative importance of species sorting versus immigration from influent wastewater on different-sized aggregates over a period of 6 months in a full-scale AGS WWTP in the Netherlands (Ali et al., 2019), ii) evaluating granule formation, microbial succession, and process performance of a full-scale AGS WWTP in Sweden over a period of 490 days (Ekholm et al., 2022), iii) comparing the mixed liquor microbial community structure and function between a full-scale AGS WWTP and a parallel full-scale AS WWTP in Sweden (Ekholm et al., 2024), and iv) evaluating different omic approaches for characterizing AGS granules (~2.0 mm) microbiome in three full-scale AGS WWTPs in the Netherlands (Kleikamp et al., 2023). Based on the literature, it could be hypothesized that each aggregate size in AGS enables the selection of a specific core microbial community that is both abundant and preserved. However, the validity of this hypothesis remains uncertain, particularly across multiple geographical distinct full-scale AGS WWTPs. This study aimed to identify the core microbial community within different-sized aggregates in full-scale AGS WWTPs undergoing stable performance in terms of carbon and nutrient removal. This contrasts with previous studies that focused on core analysis in geographically distributed full-scale activated sludge (AS) WWTPs (Chen et al., 2020; Lin et al., 2023; Saunders et al., 2016; Vestergaard et al., 2024). Particularly, this study aims to provide a comprehensive understanding of microbial beta diversity between different-sized aggregates, identify core microbial communities in different-sized aggregates, elucidate the interactions and functional roles of core microbial communities, and evaluate their relevance to wastewater treatment performance. This information is crucial for AGS WWTP operators to develop strategies to enhance system performance.

2. Materials and methods

2.1. Selection of the full-scale AGS WWTPs to sample

The sampling protocol was developed in collaboration with Nereda® (Nereda® is a trademark of Royal HaskoningDHV), who facilitated the contact and coordination for sampling. The original plan was to sample full-scale AGS WWTPs from several continents, however, due to regulatory constraints in some countries, we focused the sampling campaign only on Europe. Currently, up to 63% of full-scale AGS WWTPs worldwide are in Europe. Samples were collected from nine Nereda® AGS WWTPs (Dungannon, Cork, Midlands, Utrecht, Swindon, Clon, Manchester, Inverurie, and Dodewaard), treating municipal wastewater, situated across five European countries (England, Netherlands, Scotland, Northern Ireland, and Ireland). Samples were collected during the summer (June and July) where the operating temperature ranged between 17 °C to 18.5 °C. The nine full-scale AGS WWTPs (Table S1) were selected based on their consistently high performance in terms of carbon and nutrient removal over the past two years, achieving average removal rates of 96% for biochemical oxygen demand (BOD), 95% for COD, 96% for nitrogen (N), and 86% for phosphorus (P) (Fig. S1). These criteria were essential for identifying core organisms involved in stable and effective COD and nutrient (N and P) removal.

2.2. Collection and processing of mixed-liquor samples

For each full-scale AGS WWTP, triplicate influent and mixed liquor samples (400 mL each) were collected. The mixed liquor was collected at a depth of 1 meter below the surface and 30 min after the start of aeration to ensure the samples were uniformly mixed. The samples were then filtered using a series of cell strainers with pore sizes of 1 mm, 0.5 mm, and 0.2 mm (pluriSelect Life Science, Germany) to obtain triplicate samples for each biomass fraction: flocs (< 0.2 mm), small granules (0.2–0.5 mm), medium granules (0.5–1 mm), and large granules (> 1 mm). The classification of granule sizes was based on the criteria outlined in a previous study analyzing different granule sizes in a full-scale AGS WWTP (Ali et al., 2019). Each filter was washed with 20 mL sterile sodium phosphate buffer after processing each sample. Across the sampled WWTPs, the mixed liquor (dry biomass) was composed of 13% flocs, 22% small and medium granules, and 65% large granules (Fig. S1). All collected samples ($n = 135$) were immediately stored at −80 °C to preserve their quality until DNA extraction.

2.3. DNA extraction, amplicon sequencing library preparation, and nanopore sequencing

DNA extraction was conducted using a modified FastDNA Spin kit for Soil protocol (MP Biomedicals, USA). For each sample, 500 µL of the sample, 480 µL sodium phosphate buffer, and 120 µL MT buffer were combined in a lysing matrix E tube. Subsequently, bead beating was performed at 6 m/s for 4 cycles of 40 s each (Albertsen et al., 2015). Gel electrophoresis using Tapestation 2200 and genomic DNA screen-tapes (Agilent, USA) validated DNA product size and purity. Lastly, DNA concentration was quantified using the Qubit dsDNA HS/BR Assay kit (Thermo Fisher Scientific, USA). The extracted DNA from each sample type (i.e., influent, flocs, small granules, medium granules, and large granules) for each WWTP were pooled before sequencing ($n = 45$).

Amplicon libraries for the bacterial 16S rRNA gene variable regions 1 to 8 (V1–8) were prepared using a custom protocol. Up to 25 ng of extracted DNA served as a template for polymerase chain reaction (PCR) amplification. Each PCR reaction (50 µL) contained 0.2 mM dNTP mix, 0.01 units of Platinum SuperFi DNA Polymerase (Thermo Fisher Scientific, USA), and 500 nM of each forward and reverse primer in the supplied SuperFI Buffer. The PCR was performed with the following program: initial denaturation at 98 °C for 3 min, 25 cycles of amplification (98 °C for 30 s, 62 °C for 20 s, 72 °C for 2 min), and a final

elongation at 72 °C for 5 min. Custom 24 nt barcode sequences were used, followed by the primer sequences targeting V1-V8. The forward and reverse primers were as follows: [8F] AGRGTTYGATYMTGGCTCAG and [1391R] GACGGGCGGTGWGTRCA (Cole et al., 2009; Karst et al., 2021; Klindworth et al., 2013). The resulting amplicon libraries underwent purification using the standard protocol for CleanNGS SPRI beads (CleanNA, NL) with a bead to sample ratio of 3:5. The DNA was eluted in 25 μL of nuclease-free water (Qiagen, Germany). Sequencing libraries were prepared from the purified amplicon libraries using the SQKLSK114 kit (Oxford Nanopore Technologies, UK) according to the manufacturer's protocol with the following modifications: 500 ng total DNA was used as input, and CleanNGS SPRI beads were used for library cleanup steps. The DNA concentration was measured using the Qubit dsDNA HS Assay kit (Thermo Fisher Scientific, USA). Gel electrophoresis with Tapestation 2200 and D1000/High sensitivity D1000 screentapes (Agilent, USA) validated product size and purity of a subset of amplicon libraries.

The resulting sequencing library was loaded onto a MinION R10.4.1 flow cell and sequenced using MinKNOW 23.04.6 software (Oxford Nanopore Technologies, UK). Reads were base called and demultiplexed using MinKNOW guppy g6.5.7 with the super accurate base calling algorithm (config r10.4.1_400bps_sup.cfg) and custom barcodes.

All sequencing data associated with this project are available at the National Center for Biotechnology Information (NCBI) under BioProject PRJNA1149575. The raw sequencing reads can be accessed via the Sequence Read Archive (SRA) with accession numbers provided in Table S2.

2.4. Bioinformatic processing

The sequencing reads in the demultiplexed and basecalled fastq files were filtered for length (320–2000 bp) and quality (phred score > 15) using a local implementation of filtlong v0.2.1 with the settings –min_length 320 –max_length 2000 –min_mean_q 97. The SILVA 16S/18S rRNA 138 SSURef NR99 full-length database in RESCRIPt format was utilized from QIIME (Quast et al., 2012; Robeson et al., 2020; Yilmaz et al., 2014). Potential generic placeholders and dead-end taxonomic entries were cleared from the taxonomy flat file. Entries containing terms like 'uncultured', 'metagenome', or 'unassigned' were replaced with blank entries. The filtered reads were mapped to the SILVA 138.1 99% NR database using minimap2 v2.24r1122 with the ax mapont command (Li, 2018), followed by downstream processing with samtools v1.14 (Danecek et al., 2021). Mapping results were filtered to ensure that the query sequence length relative to alignment length deviated by less than 5%. Low-abundance operational taxonomic units (OTUs) making up less than 0.01% of the total mapped reads within each sample were disregarded as a data denoising step.

2.5. Microbial community analyses

Taxonomic beta-diversity between samples based on country, WWTP, and sample type (i.e., influent, flocs, small granules, medium granules, and large granules) were calculated using the Bray–Curtis (abundance-based) metric with the vegdist function from the vegan R package version 2.6–4 (Oksanen et al., 2022). These distances were visualized using principal component analysis (PCA) with the ampvis2.1 R package (Andersen et al., 2018). A heatmap analysis was conducted to visualize the abundance of the top 35 most abundant functional genera using the amp_heatmap function with normalize = TRUE and min_abundance = 0.1 parameters from the ampvis2.1 R- package (Andersen et al., 2018). Core OTUs and core genera were determined as taxa present with relative abundance > 0.1% and detected in over 80% of the nine sampled WWTPs using the ampvis2 package (Andersen et al., 2018). Network-based analysis and visualization were performed with open-source software Cytoscape v.3.5.148. One-way analysis of variance (ANOVA) was performed using MetaboAnalyst 6.0 (Pang et al., 2024), and a significance level was considered with p-values < 0.05. The random forest model parameters included the construction of 1000 trees with the 'random seed' set to ensure reproducibility.

3. Results and discussion

3.1. Different AGS WWTPs have different microbial community composition

In total, 45 DNA samples representing influent wastewater and various aggregate sizes (i.e., flocs, small granules, medium granules, and large granules) from the nine full-scale AGS WWTPs across Europe were sequenced, resulting in 3,414,648 non-chimeric, quality-filtered reads (Table S3). The microbial community composition of each sample was assessed using barcoded amplicons clustered into OTUs at a 97% similarity threshold (Saunders et al., 2016). Through this analysis, a total of 6110 OTUs were identified among all the analyzed samples. Given the significant diversity of reads and OTUs, the initial analysis aimed to conduct a beta diversity analysis of the nine full-sale AGS WWTPs using PCA to uncover the underlying patterns within the microbial communities (Fig. 1 and Fig. S2). It was observed that WWTPs from the same country (e.g., Netherlands or Ireland) did not cluster together; instead, WWTP-specific clustering was evident (Fig. 1a). This suggests that geographical proximity does not dictate microbial community similarity. These findings align with previous full-scale AS WWTP research based on activated sludge, which underscores the critical role of operational parameters over location in shaping microbial community structure (Matar et al., 2017; Vestergaard et al., 2024). When including the influent wastewater samples in the PCA analysis, it was observed that the influent wastewater samples are distinctly separated from the mixed liquor components (i.e., flocs, small granules, medium granules, and large granules) (Fig. 1b). The clear separation of influent wastewater samples from the mixed liquor components (Fig. 1b) suggests that species sorting induced by the AGS operational parameters and process control shapes enrichment patterns by favoring specific microbial communities. This supports the idea that immigration introduces diverse species that enhance functional diversity, while species sorting promotes competitive subpopulations based on nutrient availability and operational conditions (Leibold et al., 2004).

Despite the high selective pressure exerted by AGS, the ecological dynamics of microbial communities in AGS systems varied significantly across aggregate sizes. For instance, smaller aggregates (<1 mm), such as flocs and small to medium granules, exhibited a closer relationship with the influent microbiome compared to larger granules (Fig. 1b), highlighting the increased susceptibility of smaller aggregates to immigration from influent wastewater. In contrast, large granules tended to cluster together and far from the smaller aggregates regardless of the WWTP (Fig. 1c and Fig. S2). The observed dynamics align with previous studies showing that species sorting plays a dominant role in shaping larger aggregates, while smaller aggregates (i.e., flocs and small to medium granules) are more influenced by immigration (Ali et al., 2019). Enhanced settling of large granules allows them to utilize the readily biodegradable COD during feeding more than smaller aggregates, reducing competition for resources and leading to a more selective community (Ali et al., 2019; Hamiruddin et al., 2021). Furthermore, their extended SRT provides more time for microbial communities to mature and stabilize, making large granules more resistant to the influence of new immigrants from the influent (Ali et al., 2019; Hamiruddin et al., 2021). Collectively, these factors lead to a more selective and similar community in large granules across different full-scale AGS WWTPs.

3.2. Small subset of highly shared and abundant OTUs observed across the nine full-scale AGS WWTPs

The PCA results revealed WWTP-specific clustering patterns (Fig. 1

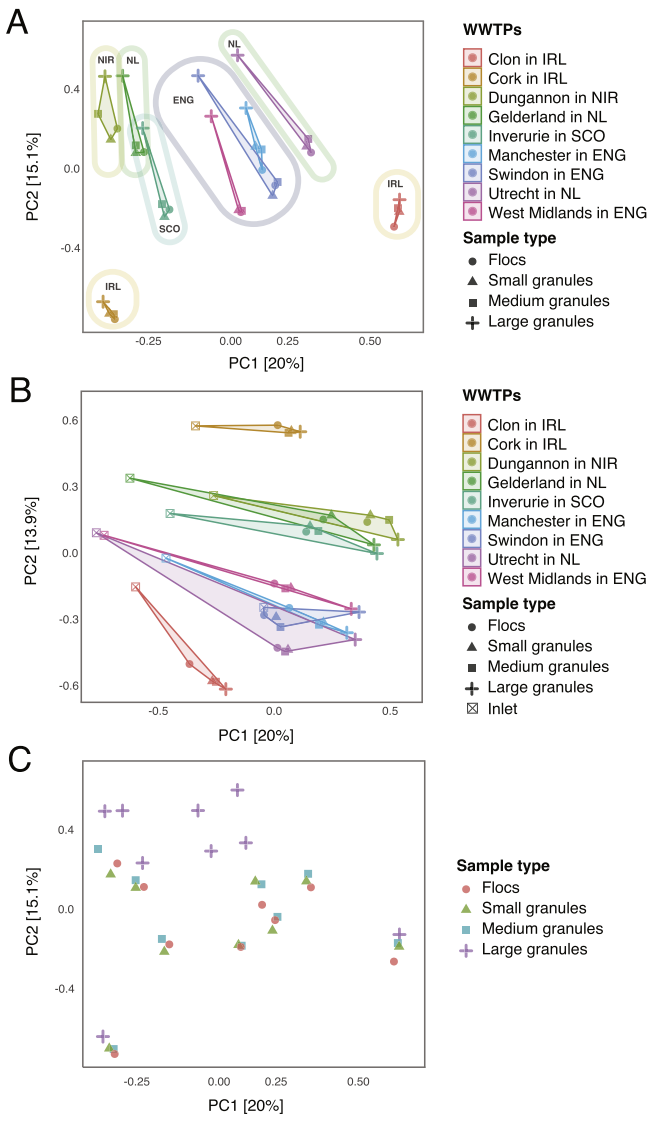


Fig. 1. Principal component analysis (PCA) revealing the beta diversity between the nine full-scale AGS WWTPs based on: A) mixed liquor components (flocs, small granules, medium granules, and large granules) in the nine WWTPs, B) mixed liquor components and inlet (i.e., influent wastewater) in the nine WWTPs, and C) mixed liquor components. IRL stands for Ireland, NIR for Northern Ireland, SCO for Scotland, ENG for England, and NL for the Netherlands.

and Fig. S2). Therefore, the next objective is to identify the OTUs consistently present across different WWTPs and sample types (i.e., flocs, small granules, medium granules, large granules). The premise for this analysis is that despite WWTP-specific clustering patterns, the shared OTUs within the same sample type (i.e., flocs, small granules, medium granules, and large granules) across the different WWTPs are primarily selected by the AGS operational parameters and process control, thereby these shared OTUs play a crucial role in system performance and stability. To address this, the frequencies of shared OTUs within each sample type and across the nine AGS WWTPs were examined (Fig. 2A). Additionally, the relative abundance of these shared OTUs was analyzed to understand their significance within the system, with OTUs showing higher relative abundance are expected to exert a greater influence on treatment processes (Fig. 2b) (Saunders et al., 2016).

The frequency analysis revealed a skewed distribution in the occurrence of each OTU across all sample types, with most OTUs predominantly present in a single WWTP. In contrast, the number of shared OTUs progressively declined as the number of WWTPs increased, with the fewest shared OTUs observed across all nine plants (Fig. 2b). Interestingly, the analysis of the relative abundance patterns of the shared OTUs revealed an opposite trend compared to the frequency analysis, where the relative abundance of shared OTUs increased with the increase in the number of WWTPs (Fig. 2b). These analyses reveal that each size aggregate contains a large number (1000–2000) of low-abundance OTUs that are unique to a specific WWTP. Previous studies on full-scale WWTPs based on activated sludge have observed similar patterns, with numerous unique low-abundance OTUs specific to each WWTP (Matar et al., 2017; Saunders et al., 2016). Research in AGS, AS, soil, and groundwater have shown that low-abundance OTUs are predominantly shaped by stochastic factors, with dispersal playing a key role in their distribution (Ali et al., 2019; Kim et al., 2013; Richter--Heitmann et al., 2020; Zhong et al., 2023). Consequently, they may be more vulnerable to mass effects and neutral processes, making them prone to changes over time (Lindström and Langenheder, 2012; Richter-Heitmann et al., 2020; Zhong et al., 2023). These results suggest that variability in microbial community composition across the different full-scale AGS WWTPs (Fig. 1) may be largely driven by these unique low-abundance OTUs.

It has been reported that low-abundance OTUs in full-scale WWTPs based on AS have minimal impact on treatment performance, instead acting as microbial 'fingerprints' that differentiate between facilities (Matar et al., 2017; Saunders et al., 2016). In contrast, highly abundant OTUs in AS WWTPs have been linked to a selective advantage conferred by the treatment process, making them essential for sustaining system functionality and resilience in wastewater treatment (Saunders et al., 2016). In the current study, a small subset of highly abundant OTUs was shared by 80% (representing a relative abundance of approximately 17%) to 100% (representing a relative abundance of approximately 30%) of the nine AGS WWTPs, indicating that these OTUs were highly selected by AGS operational parameters and process control. Therefore, focusing on the shared and abundant OTUs in AGS WWTPs provide a clearer understanding of the true ecological dynamics among different-sized aggregates, rather than examining the entire community.

3.3. *Core OTUs are consistently selected in the different-sized aggregates*

Shared OTU analysis underscored the relevance of abundant and prevalent OTUs within each size aggregate. Therefore, the next step was to characterize the core OTUs, meaning those that are abundant (>0.1% of relative abundance) and present in at least 80% of the nine full-scale AGS WWTPs, as reported in previous studies analyzing full-scale WWTPs (Dueholm et al., 2022). The analysis of OTUs across the different-sized aggregates revealed species richness ranging from 2559 to 3586 OTUs, with core OTUs ranging between 128 and 139 OTUs and accounting for 50% to 53% of the total relative abundance (Figs. 3a and b). Specifically, large granules contained 137 core OTUs, small and medium granules had 139, and flocs contained 128 core OTUs, suggesting a stable core microbiome persists across varying aggregate sizes. Furthermore, the taxonomic distribution of these core OTUs (Fig. 3c) varied across samples, encompassing a spectrum of 8 to 11 phyla, 15 to 21 classes, 28 to 35 orders, 44 to 50 families, and 54 to 60 genera, with a total of 84 distinct genera identified. Notably, the genus level exhibited the greatest diversity, with most core OTUs successfully taxonomically classified. However, the proportion of classified genera was only 73% in large granules compared to 82% in flocs and medium granules and 81% in small granules, suggesting that larger aggregates host a less well-defined taxon.

A comprehensive analysis of the core OTUs in the different-sized aggregates, followed by dereplication, identified 229 unique OTUs, representing 4% of the total observed OTUs. The 229 OTUs were consistently detected across the different-sized aggregates (Fig. S3), indicating a significant overlap in microbial composition among the size

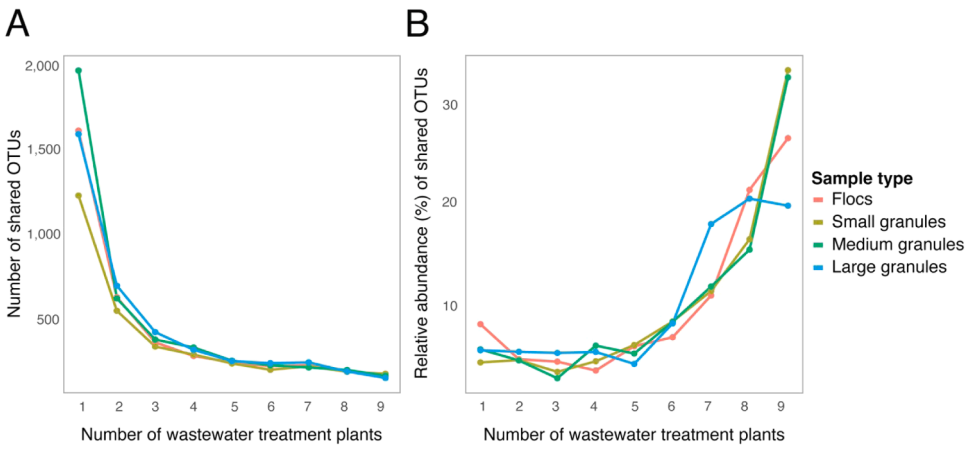


Fig. 2. Frequency and relative abundance of OTUs across influent, flocs, small granules (SG), medium granules (MG), and large granules (LG) in the nine full-scale AGS WWTPs. A) Distribution of OTUs among different WWTPs for each sample type. In other words, it displays the number of OTUs exclusive to individual WWTPs, those shared among two to eight WWTPs, and those shared among all plants (i.e., nine plants). B) Relative abundance of OTUs identified in Fig. 2a across different sample types throughout the nine full-scale AGS WWTPs.

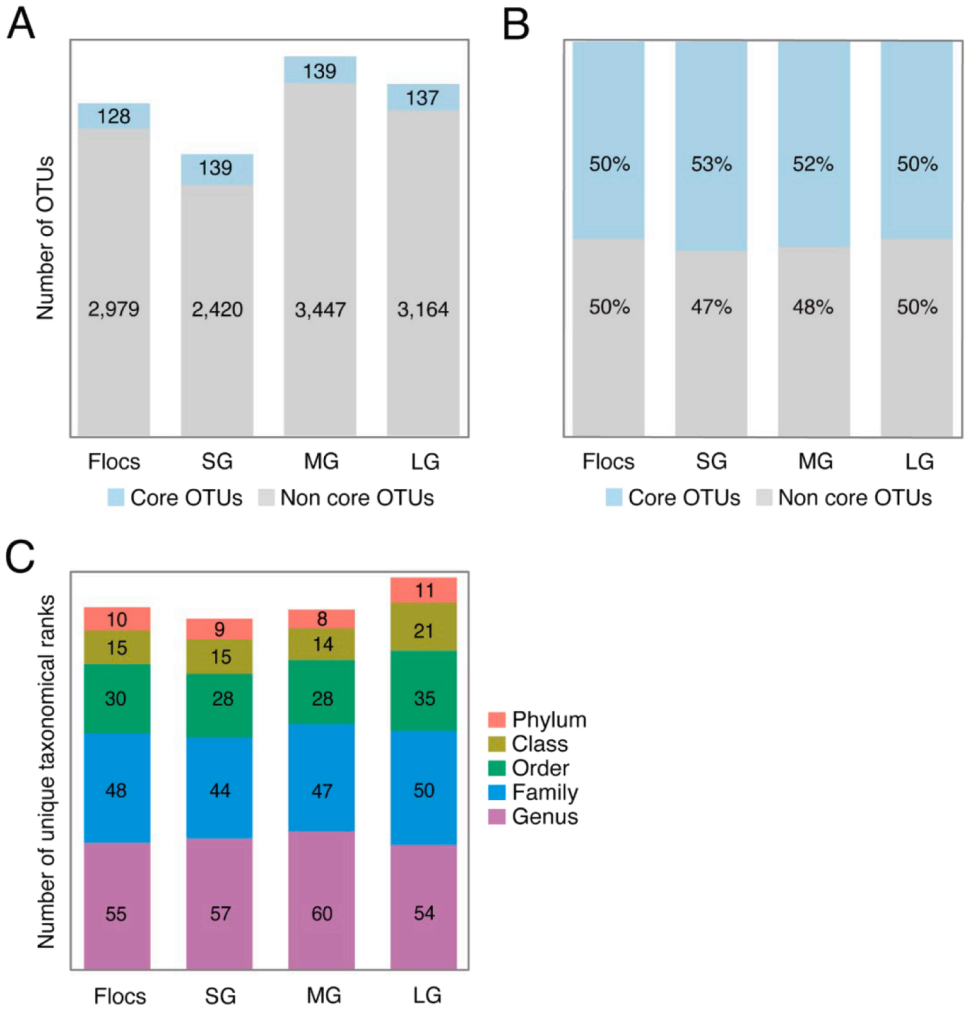


Fig. 3. Analysis of core and non-core OTUs in the different-sized aggregates. A) Number of core OTUs vs. non-core OTUs. B) Relative abundance of core and non-core OTUs. C) Number of core phylum, class, order, family, and genus. SG stands for small granules, MG for medium granules, and LG for large granules.

aggregates. To investigate the extent of this overlap, ANOVA was employed to discern any distinct enrichment patterns across the different-sized aggregates (Fig. 4). Notably, 212 OTUs showed no significant differences in abundance between aggregate sizes, suggesting their selection is likely driven by system-wide selective pressures rather than aggregate-specific factors. In contrast, 17 out of the 229 OTUs displayed statistically significant differences ($p < 0.05$) in relative abundance across the different-sized aggregates. These OTUs exhibiting

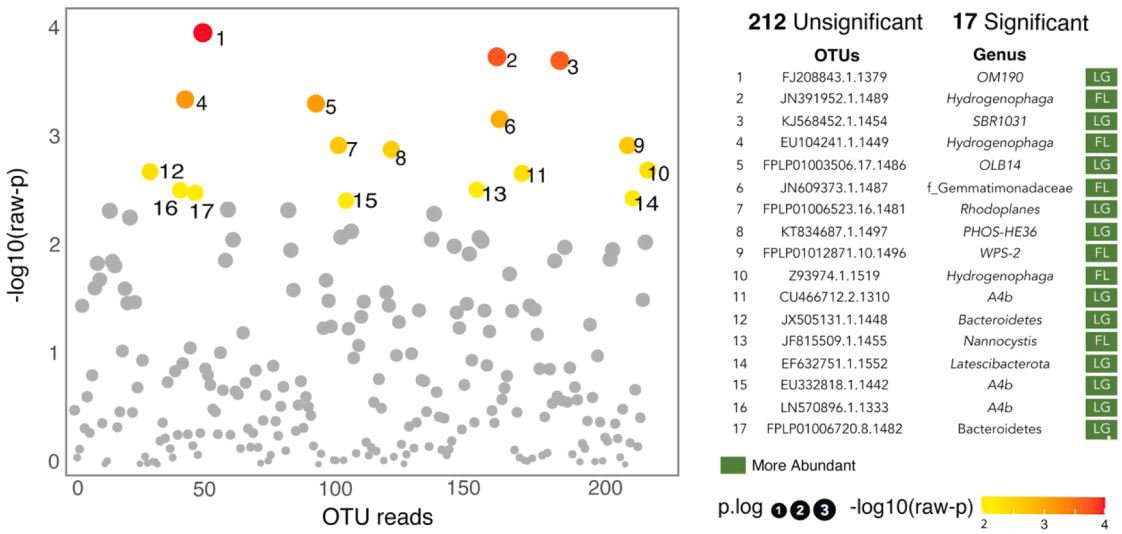


Fig. 4. ANOVA analysis of core OTU reads in different-sized aggregates from the nine full-scale AGS WWTPs. This figure presents the results of an ANOVA analysis based on core OTU reads from different-size aggregates in the nine full-scale AGS WWTPs. Each dot represents an OTU, and the size of the dot represents the statistical significance (p-value) on a log scale. Gray dots represent 212 OTUs that did not show significant differences among size aggregates ($p > 0.05$), while colored dots represent 17 OTUs with significant differences ($p \leq 0.05$). The color gradient ranges from yellow (indicating lower differences) to red (indicating higher differences). The accompanying table presents the OTU IDs, genus classifications, and aggregate sizes with significant differences, where FL represents flocs and LG represents large granules. The green color indicates higher OTU abundance in the specified size aggregate.

significant differences were primarily enriched in larger granules or flocs, pointing to distinct microbial communities within these aggregates (Fig. 4). Furthermore, a random forest analysis, a supervised machine learning method that constructs multiple decision trees based on input data with known classifications, was utilized to predict the size of aggregates based on core OTU profiles (Breiman, 2001). The model achieved an overall accuracy of 40%. The effectiveness in prediction varied across aggregate sizes, reaching 78% for large granules and 67% accuracy for flocs, yet there was a 0% success rate in distinguishing between small and medium granules. These findings indicate that despite extensive microbial sharing and similarity in enrichment patterns, subtle differences still allow for the differentiation of large granules and, to a lesser extent, flocs. This is consistent with previous microbial composition analyses, where flocs and large granules are distinct, while small and medium granules cluster together (Fig. 1).

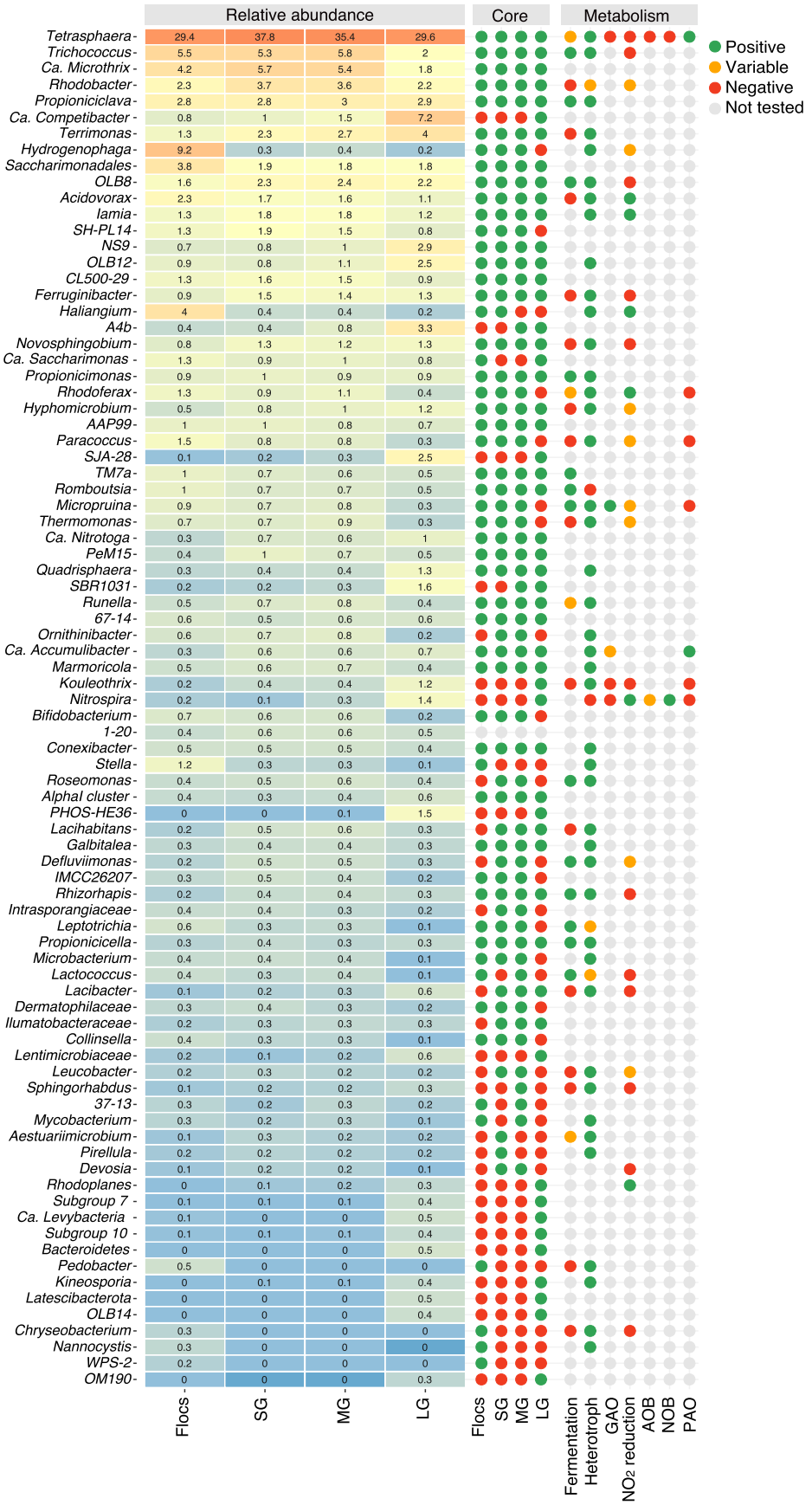
3.4. *Core genera show unique prevalence patterns distinct to each size aggregate*

The OTU analysis showed the presence of core OTUs across all aggregate sizes, though the prevalence or abundance of some OTUs differs based on aggregate size. These findings emphasize the need to investigate size-specific core genera to understand where certain core genera are most likely to prevail. To achieve this, a heatmap was employed to illustrate the mean relative abundances of the core genera, indicating the presence or absence of each core genus in the different-sized aggregates (Fig. 5). Each genus was annotated with potential functions according to the MiDAS Field Guide (Fig. 5). A co-occurrence network analysis was employed to visualize interactions among the core genera of each size aggregate, presenting their potential functions through nodes' color and labels (Fig. 6). A Venn diagram was also constructed to illustrate the shared genera patterns among the different-sized aggregates (Fig. 6). This approach aims to provide a comprehensive view of both the distribution and potential ecological roles of core genera across the various aggregate sizes.

Collectively, these analyses revealed that although each size aggregate comprises between 54 and 60 core genera, 33 are consistently shared across all aggregate sizes (Fig. 5 and 6). This extensive sharing underscores the widespread presence of the core microbial community across different aggregate sizes, corroborating previous observations (Fig. 4) that full-scale AGS WWTPs selectively maintain specific core microbial community regardless of aggregate size. More importantly, the shared core genera include functional groups such as fermenters, aerobic heterotrophs, nitrite reducers, nitrifying bacteria (*Ca.* Nitrotoga), and PAOs. The shared PAO genera feature *Ca.* Accumulibacter and *Tetrasphaera*-related PAOs. However, due to the limitations of 16S rRNA gene in accurately classifying *Tetrasphaera*-related PAOs, this group could also potentially encompass the genera *Knoellia, Ca.* Lutibacillus, and *Ca.* Phosphoribacter (Ruiz-Haddad et al., 2024; Singleton et al., 2022). Furthermore, the remaining OTUs that are not shared among all aggregates demonstrated unique patterns distinct to each size aggregate. Within large granules, 15 unique genera were identified, including the GAO *Ca.* Competibacter and the nitrifying bacteria *Nitrobacter*. However, the functional potential of most other genera unique to large granules remains uncharacterized. Additionally, large granules shared a small subset of genera with other aggregates, mainly with small and medium granules, and these shared genera were predominantly involved in carbon removal. Conversely, flocs harbored five distinct genera, predominantly aerobic heterotrophs. Most of the other genera found in flocs primarily include nitrite reducers and the GAO *Micropruina*, which are also shared with small and medium granules.

The above results suggest that the functional groups responsible for N and P removal are predominantly shared across all aggregates (Figs. 5 and 6). Moreover, the complete functional potential of most core genera (31 out of 84) relevant to AGS system, especially in larger aggregates, remains largely unexplored (Figs. 5 and 6). This highlights the need for further research to fully understand their ecological roles. Metagenome-assembled genomes (MAGs) provide a powerful method for uncovering the functional capabilities of previously uncharacterized genera, which is particularly relevant in complex sludge microbial communities (Ruiz-Haddad et al., 2024). For instance, in AGS systems, MAGs have facilitated the discovery of novel PAOs within previously unidentified genera by revealing their potential for polyphosphate accumulation, thereby enabling targeted experiments to validate their phenotypes (Singleton et al., 2022).

Previous studies have demonstrated that the presence and relative abundance of core organisms are critical in maintaining the stability and performance of wastewater treatment processes (Lin et al., 2023;



(caption on next page)

Fig. 5. Heatmap of the relative abundances of core genera across the different aggregate sizes. The heatmap is complemented with information on the presence or absence of core genera in the different aggregate sizes. Additional information on functional potential is included, with data from the MiDAS database for carbon-related metabolism, nitrification, and denitrification. Key functional groups such as glycogen-accumulating organisms (GAOs), ammonia-oxidizing bacteria (AOB), nitrite-oxidizing bacteria (NOB), and polyphosphate-accumulating organisms (PAOs) are also highlighted. SG stands for small granules, MG for medium granules, and LG for large granules.

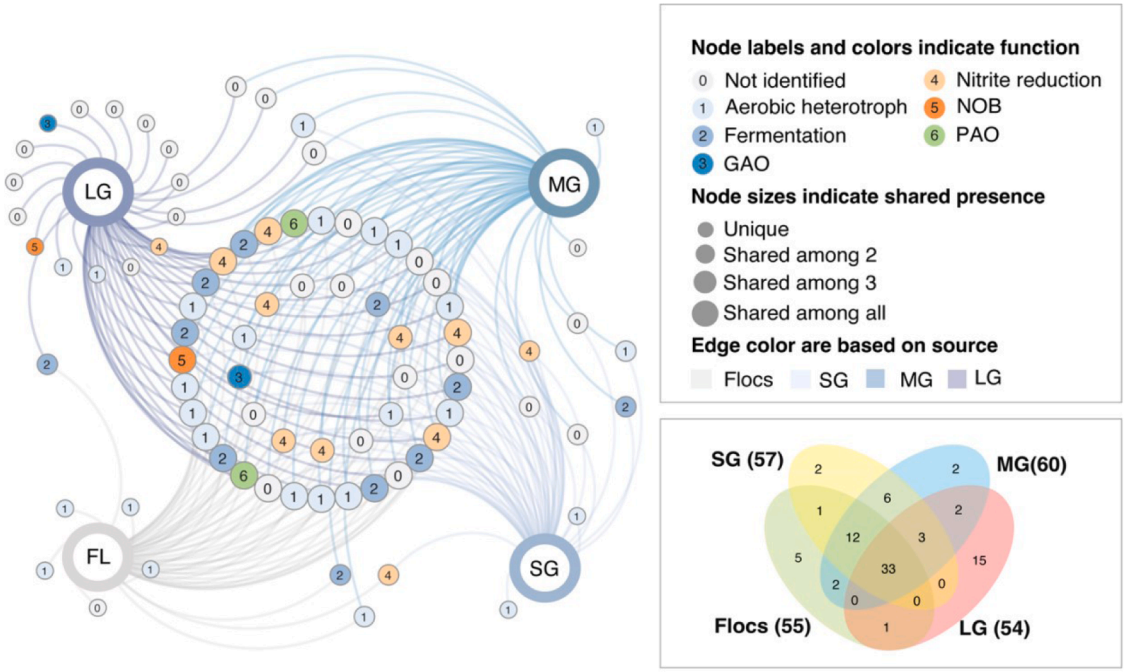


Fig. 6. Microbial co-occurrence analysis of core genera across aggregate sizes. Nodes represent genera and are color-coded by their reported functions. The node size reflects the degree of uniqueness of genera to specific aggregate sizes, with the largest node size indicating genera common across all aggregate sizes. Edge colors denote the source of the genera: flocs (FL), small granules (SG), medium granules (MG), and large granules (LG). Additionally, the Venn diagram summarizes the shared genera among the four aggregate sizes observed in the genera-network analysis. The numbers in brackets represent the number of classified core genera in each aggregate size.

Nielsen et al., 2010; Saunders et al., 2016; Świątczak and Cydzik-Kwiatkowska, 2018). The current study, along with previous studies, shows that these core communities are enriched with key functional groups, such as PAOs, GAOs, and nitrifiers (Lin et al., 2023; Nielsen et al., 2010; Saunders et al., 2016; Świątczak and Cydzik-Kwiatkowska, 2018). Additionally, artificial neural network models in AS systems have identified the presence and relative abundance of these functional groups as reliable biomarkers for predicting WWTP performance (Liu et al., 2023). Therefore, the core genera identified in our study could serve as biomarkers for assessing and predicting WWTP stability and performance. To validate this concept, future research should focus on larger datasets that compare the core microbial communities in well-performing and poorly-performing AGS WWTPs. Such studies would be essential for enabling WWTPs operators to monitor community dynamics effectively and devise strategies to maintain optimal treatment performance. This can be achieved through on-site 16S rRNA gene amplicon sequencing using platforms such as the Oxford Nanopore MinION, which provides a portable and real-time monitoring of microbial community dynamics (Latorre-Pérez et al., 2020; Leggett et al., 2019). This technology enables a detailed analysis of changes in the composition and abundance of core organisms, offering more accurate predictions of system stability and performance shifts.

3.5. Key functional groups are enriched in specific aggregate sizes

The analysis aimed to elucidate the enrichment patterns of key functional groups across different-sized aggregates (Fig. 7). To achieve this, the data was divided into two components: total genera, comprising all functional genera (i.e., both core and non-core) identified across samples, and core genera, including only the genera belonging to the core community. The relative abundance of each genus within a functional group was then summed to determine the total relative abundance of each functional group in both the total and core genera, allowing for a comparison of relative abundance and enrichment patterns between them. The analysis focused on three functional groups: PAOs (e.g., *Ca.* Accumulibacter, *Dechloromonas, Microlunatus*, and *Tetrasphaera*-related PAOs), GAOs (e.g., *Defluviicoccus, Ca.* Competibacter, *Micropruina*, and *Contendobacter*), and nitrifiers (AOB such as *Nitrosospira* and *Nitrosomonas*, and NOB such as *Nitrospira, Nitrotoga*, and *Nitrobacter*) (Dueholm et al., 2022; Ruiz-Haddad et al., 2024). It is important to note that these findings may not be generalizable to all nine AGS WWTPs, as the microbial composition within each size aggregate has been observed to vary across different WWTPs (Fig. 1C). Nevertheless, this approach can still uncover functional group preferences specific to certain aggregate sizes.

The relative abundances of functional groups, including PAOs, GAOs, and nitrifiers, exhibited consistent patterns across different-sized aggregates, irrespective of whether the analysis focused solely on core genera or included the entire community (i.e., core and non-core) (Fig. 7B-E). Most of the relative abundances within the functional groups were attributed to the core genera, while non-core functional genera contributed only minimally. Consistent with previous findings (Fig. 6), all functional groups were present across all size aggregates, though certain functional groups exhibited higher relative abundances in specific size aggregates. Most relative abundances were attributed to core genera, while non-core functional genera contributed minimally.

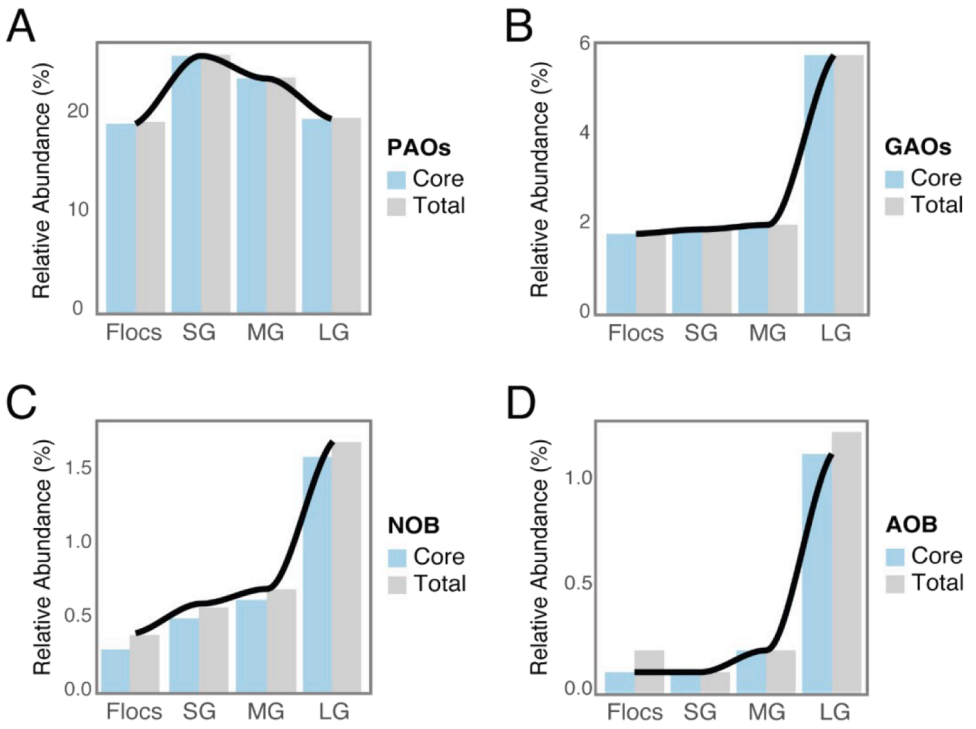


Fig. 7. Functional microbial composition across the different-sized aggregates. A-D). Graphs depicting the relative abundances of the functional groups across the different-sized aggregates: A) PAOs, B) GAOs, C) NOB, and D) AOB. Each graph features two columns representing the analyzed groups: Total (gray), which encompasses all functional genera (i.e., both core and non-core), and Core (light blue), which shows the relative abundances of functional genera found exclusively in the core.

Tetrasphaera-related PAOs were the most abundant functional group in all size aggregates, with the highest enrichment observed in granules compared to flocs. Ali et al. (2019) reported an even distribution of *Tetrasphaera* across different granule sizes and a significant enrichment of PAOs, primarily *Ca.* Accumulibacter, in larger granules. However, findings in the current study show higher enrichment of PAOs in small granules compared to medium and large granules. This increased abundance of PAOs in smaller granules (0.2–0.6 mm) could be explained by their higher surface-to-volume ratio, which enhances aerobic phosphorus removal and reduces competition with GAOs (Nguyen Quoc et al., 2021a). Additionally, the *Tetrasphaera* dominance over *Ca.* Accumulibacter in the current study may be attributed to the ability of *Tetrasphaera* to utilize a broader range of substrates (Herbst et al., 2019; Singleton et al., 2022).

The functional groups GAOs, NOB, and AOB were predominantly enriched in large granules, consistent with the presence of core genera such as *Nitrospira, Nitrosomonas*, and *Ca.* Competibacter. This enrichment aligns with previous studies and is likely driven by the longer SRT in large granules, which support the growth of these slow-growing organisms. Additionally, oxygen gradients have been shown to promote synergistic interactions within granules, with AOB proliferation occurring in the surface layers, where highly aerobic conditions prevail, while NOB and GAOs are favored in the deeper, more anoxic layers (Guimarães et al., 2017; Layer et al., 2019; Nguyen Quoc et al., 2021a, 2021b; Weissbrodt et al., 2013). Furthermore, the increased substrate availability in large granules may benefit these genera, which have limited substrate uptake capacity, particularly as they rely on diffusible substrates such as volatile fatty acids (VFAs), ammonia, and nitrite that can permeate the biofilm (Cai et al., 2018; Guimarães et al., 2023; Layer et al., 2019; McIlroy et al., 2014).

A genus-specific selection pattern was observed for GAOs, with *Micropruina* being more prevalent in smaller aggregates (< 1 mm), while *Ca.* Competibacter predominated in larger granules. This pattern may be due to the preference of *Ca.* Competibacter for VFAs, which are more concentrated at the bottom of the reactor and can easily diffuse through larger aggregates (Ali et al., 2019; Guimarães et al., 2023; Layer et al., 2019). In contrast, *Micropruina*, recognized for its ability to utilize a wide range of substrates, particularly carbohydrates hydrolyzed from polymers, likely benefits from the increased surface-to-volume ratio of smaller aggregates (McIlroy et al., 2018; Nguyen Quoc et al., 2021a). Furthermore, the production of fermentative products by *Micropruina* has been reported to provide substrate for PAOs.

Flocs did not exhibit the highest abundance of any functional group. They displayed similar relative abundances of AOB, GAOs, and NOB as to small and medium granules, and a similar relative abundance of PAOs as to large granules. The similarities in enrichment patterns between flocs and other aggregates suggest that significant abrasion from larger aggregates into floccular structures may be occurring, especially since flocs are not expected to access soluble substrates during the feeding phase (Van Dijk et al., 2022).

Collectively, these results reveal similar enrichment patterns among certain aggregates, indicating the role of AGS parameters and process control in selecting for functional groups across all aggregate sizes, though specific functional groups consistently remain predominantly associated with a particular size aggregate. For instance, aerobic heterotrophs are more abundant in flocs, PAOs in small granules, and GAOs and nitrifiers in larger granules. Despite the fact that some functional groups are enriched more in particular aggregate sizes, it is important to note that previous studies in AGS have reported discrepancies between the relative abundance and activity of different genera, such as *Nitrospira, Ca.* Competibacter, and *Ca.* Accumulibacter (Kleikamp et al., 2023; Nguyen Quoc et al., 2021a). These findings highlight that relative abundance does not always correlate with microbial activity. Therefore, future research should focus on quantifying the activity of these functional groups in different-sized aggregates using metatranscriptomics and metaproteomics (Kleikamp et al., 2023) to better understand how enrichment correlates with functional activity in specific aggregate sizes and to guide practitioners to augment poorly-performing AGS plants

with the right aggregate sizes to enhance the activity of certain functional groups.

4. Conclusions

- Beta-diversity analysis showed that variability in microbial community composition among aggregates across different WWTPs was largely driven by WWTP-specific low-abundance OTUs, whose presence is possibly driven by stochastic processes. In contrast, each aggregate size consistently harbors a subset of prevalent and abundant OTUs.
- 229 OTUs were consistently detected across the different-sized aggregates due to the high selective pressure imposed by the AGS operational parameters and process control.
- All aggregates contained a small subset of 128–139 core OTUs that were both prevalent and abundant across all sizes. The core microbiome included genera associated with key functional groups responsible for fermentation, aerobic heterotrophy, nitrification, and P-removal (i.e., PAOs).
- The prevalence of functional groups varied with aggregate size, with aerobic heterotrophs dominating in flocs, PAOs being more prevalent in small granules, and GAOs and nitrifiers being more abundant in large granules.

CRediT authorship contribution statement

Lucia Ruiz-Haddad: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Dario Rangel Shaw:** Writing – original draft, Supervision. **Muhammad Ali:** Writing – review & editing, Supervision, Conceptualization. **Mario Pronk:** Writing – review & editing, Supervision, Conceptualization. **Mark C.M. van Loosdrecht:** Writing – review & editing, Supervision, Conceptualization. **Pascal E. Saikaly:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Authors of this submission, Pascal Saikaly and Mark van Loosdrecht, serve as editors of *Water Research*. The manuscript has been handled and reviewed independently by other editors of the journal, with Pascal Saikaly and Mark van Loosdrecht recusing themselves from any decisions related to the evaluation and acceptance of this work.

Acknowledgment

This work was supported by King Abdullah University of Science and Technology (KAUST).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2024.123036.

References

Albertsen, M., Karst, S.M., Ziegler, A.S., Kirkegaard, R.H., Nielsen, P.H., 2015. Back to Basics – The influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. PLoS. One 10, e0132783. https://doi.org/10.1371/journal.pone.0132783.

Ali, M., Singh, Y., Fortunato, L., Rehman, Z.U., Manjunath, S., Vrouwenvelder, J.S., Pronk, M., Van Loosdrecht, M.C.M., Saikaly, P.E., 2023. Performance evaluation of a pilot-scale aerobic granular sludge integrated with gravity-driven membrane system treating domestic wastewater. ACS EST Water 3, 2681–2690. https://doi.org/10.1021/acsestwater.3c00178.

Ali, M., Wang, Z., Salam, K.W., Hari, A.R., Pronk, M., Van Loosdrecht, M.C.M., Saikaly, P. E., 2019. Importance of species sorting and immigration on the bacterial assembly of different-sized aggregates in a full-scale aerobic granular sludge plant. Environ. Sci. Technol. 53, 8291–8301. https://doi.org/10.1021/acs.est.8b07303.

Andersen, K.S., Kirkegaard, R.H., Karst, S.M., Albertsen, M., 2018. ampvis2: an R package to analyse and visualise 16S rRNA amplicon data (preprint). Bioinformatics. https://doi.org/10.1101/299537.

Aonofriesei, F., Petrosanu, M., 2007. Activated sludge bulking episodes and dominant filamentous bacteria were observed at the wastewater treatment plant Constănţa Sud in Romania. In: Proc. Romanian Acad, B, pp. 83–87.

Begmatov, S., Dorofeev, A.G., Kadnikov, V.V., Beletsky, A.V., Pimenov, N.V., Ravin, N.V., Mardanov, A.V., 2022. The structure of microbial communities of activated sludge of large-scale wastewater treatment plants in the city of Moscow. Sci. Rep. 12, 3458. https://doi.org/10.1038/s41598-022-07132-4.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Cai, M., Ng, S.K., Lim, C.K., Lu, H., Jia, Y., Lee, P.K.H., 2018. Physiological and metagenomic characterizations of the synergistic relationships between ammonia- and nitrite-oxidizing bacteria in freshwater nitrification. Front. Microbiol. 9, 280. https://doi.org/10.3389/fmicb.2018.00280.

Chen, H., Wang, M., Chang, S., 2020. Disentangling community structure of ecological system in activated sludge: core communities, functionality, and functional redundancy. Microb. Ecol. 80, 296–308. https://doi.org/10.1007/s00248-020-01492-y.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M., Tiedje, J.M., 2009. The ribosomal database project: improved alignments and new tools for rRNA analysis. Nucleic. Acids. Res. 37, D141–D145. https://doi.org/10.1093/nar/gkn879.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10, giab008. https://doi.org/10.1093/gigascience/giab008.

De Kreuk, M.K., Heijnen, J.J., Van Loosdrecht, M.C.M., 2005. Simultaneous COD, nitrogen, and phosphate removal by aerobic granular sludge. Biotechnol. Bioeng. 90, 761–769. https://doi.org/10.1002/bit.20470.

Dueholm, M.K.D., Nierychlo, M., Andersen, K.S., Rudkjøbing, V., Knutsson, S., , MiDAS Global Consortium, Arriaga, S., Bakke, R., Boon, N., Bux, F., Christensson, M., Chua, A.S.M., Curtis, T.P., Cytryn, E., Erijman, L., Etchebehere, C., Fatta-Kassinos, D., Frigon, D., Garcia-Chaves, M.C., Gu, A.Z., Horn, H., Jenkins, D., Kreuzinger, N., Kumari, S., Lanham, A., Law, Y., Leiknes, T., Morgenroth, E., Muszyński, A., Petrovski, S., Pijuan, M., Pillai, S.B., Reis, M.A.M., Rong, Q., Rossetti, S., Seviour, R., Tooker, N., Vainio, P., Van Loosdrecht, M., Vikraman, R., Wanner, J., Weissbrodt, D., Wen, X., Zhang, T., Nielsen, Per H., Albertsen, M., Nielsen, Per Halkjær, 2022. MiDAS 4: a global catalogue of full-length 16S rRNA gene sequences and taxonomy for studies of bacterial communities in wastewater treatment plants. Nat. Commun. 13, 1908. https://doi.org/10.1038/s41467-022-29438-7.

Ekholm, J., Persson, F., De Blois, M., Modin, O., Gustavsson, D.J.I., Pronk, M., Van Loosdrecht, M.C.M., Wilén, B.M., 2024. Microbiome structure and function in parallel full-scale aerobic granular sludge and activated sludge processes. Appl. Microbiol. Biotechnol. 108, 334. https://doi.org/10.1007/s00253-024-13165-8.

Ekholm, J., Persson, F., De Blois, M., Modin, O., Pronk, M., Van Loosdrecht, M.C.M., Suarez, C., Gustavsson, D.J.I., Wilén, B.M., 2022. Full-scale aerobic granular sludge for municipal wastewater treatment – granule formation, microbial succession, and process performance. Environ. Sci. Water Res. Technol. 8, 3138–3154. https://doi.org/10.1039/D2EW00653G.

Guimarães, L.B., Gubser, N.R., Lin, Y., Zlópasa, J., Felz, S., Tomás Martínez, S., Pronk, M., Neu, T.R., Dueholm, M.K.D., Albertsen, M., Da Costa, R.H.R., Nielsen, P. H., Van Loosdrecht, M.C.M., Weissbrodt, D.G., 2023. Production of extracellular polymeric substances in granular sludge under selection for accumulibacter and competibacter (preprint). Bioengineering. https://doi.org/10.1101/2023.03.24.534144.

Guimarães, L.B., Mezzari, M.P., Daudt, G.C., Da Costa, R.H., 2017. Microbial pathways of nitrogen removal in aerobic granular sludge treating domestic wastewater: microbial pathways of nitrogen removal in granular sludge. J. Chem. Technol. Biotechnol 92, 1756–1765. https://doi.org/10.1002/jctb.5176.

Guo, J., Ni, B.J., Han, X., Chen, X., Bond, P., Peng, Y., Yuan, Z., 2017. Unraveling microbial structure and diversity of activated sludge in a full-scale simultaneous nitrogen and phosphorus removal plant using metagenomic sequencing. Enzyme Microb. Technol. 102, 16–25. https://doi.org/10.1016/j.enzmictec.2017.03.009.

Hamiruddin, N.A., Awang, N.A., Mohd Shahpudin, S.N., Zaidi, N.S., Said, M.A.M., Chaplot, B., Azamathulla, H.M., 2021. Effects of wastewater type on stability and operating conditions control strategy in relation to the formation of aerobic granular sludge – a review. Water Sci. Technol. 84, 2113–2130. https://doi.org/10.2166/wst.2021.415.

Hamza, R.A., Sheng, Z., Iorhemen, O.T., Zaghloul, M.S., Tay, J.H., 2018. Impact of food-to-microorganisms ratio on the stability of aerobic granular sludge treating high-strength organic wastewater. Water. Res. 147, 287–298. https://doi.org/10.1016/j.watres.2018.09.061.

Herbst, Dueholm, Wimmer, Nielsen, 2019. The proteome of *Tetrasphaera elongata* is adapted to changing conditions in wastewater treatment plants. Proteomes. 7, 16. https://doi.org/10.3390/proteomes7020016.

Karst, S.M., Ziels, R.M., Kirkegaard, R.H., Sørensen, E.A., McDonald, D., Zhu, Q., Knight, R., Albertsen, M., 2021. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. Nat. Methods 18, 165–169. https://doi.org/10.1038/s41592-020-01041-y.

Kim, T.S., Jeong, J.Y., Wells, G.F., Park, H.D., 2013. General and rare bacterial taxa demonstrating different temporal dynamic patterns in an activated sludge bioreactor. Appl. Microbiol. Biotechnol. 97, 1755–1765. https://doi.org/10.1007/s00253-012-4002-7.

Kim, Y.K., Yoo, K., Kim, M.S., Han, I., Lee, M., Kang, B.R., Lee, T.K., Park, J., 2019. The capacity of wastewater treatment plants drives bacterial community structure and its assembly. Sci. Rep. 9, 14809. https://doi.org/10.1038/s41598-019-50952-0.

Kleikamp, H.B.C., Grouzdev, D., Schaasberg, P., Van Valderen, R., Van Der Zwaan, R., Wijgaart, R.V.D., Lin, Y., Abbas, B., Pronk, M., Van Loosdrecht, M.C.M., Pabst, M., 2023. Metaproteomics, metagenomics and 16S rRNA sequencing provide different perspectives on the aerobic granular sludge microbiome. Water. Res. 246, 120700. https://doi.org/10.1016/j.watres.2023.120700.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., Glöckner, F.O., 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic. Acids. Res. 41, e1. https://doi.org/10.1093/nar/gks808. –e1.

Latorre-Pérez, A., Pascual, J., Porcar, M., Vilanova, C., 2020. A lab in the field: applications of real-time, in situ metagenomic sequencing. Biol. Methods Protoc. 5, bpaa016. https://doi.org/10.1093/biomethods/bpaa016.

Layer, M., Adler, A., Reynaert, E., Hernandez, A., Pagni, M., Morgenroth, E., Holliger, C., Derlon, N., 2019. Organic substrate diffusibility governs microbial community composition, nutrient removal performance and kinetics of granulation of aerobic granular sludge. Water Res. X 4, 100033. https://doi.org/10.1016/j.wroa.2019.100033.

Leggett, R.M., Alcon-Giner, C., Heavens, D., Caim, S., Brook, T.C., Kujawska, M., Martin, S., Peel, N., Acford-Palmer, H., Hoyles, L., Clarke, P., Hall, L.J., Clark, M.D., 2019. Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. Nat. Microbiol. 5, 430–442. https://doi.org/10.1038/s41564-019-0626-z.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M., Gonzalez, A., 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecol. Lett. 7, 601–613. https://doi.org/10.1111/j.1461-0248.2004.00608.x.

Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 34, 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

Lin, H., Zheng, Y., Yang, Y., Liu, F., Yang, K., Zhang, B., Wen, X., 2023. The role of the core microorganisms in the microbial interactions in activated sludge. Environ. Res. 235, 116660. https://doi.org/10.1016/j.envres.2023.116660.

Lindström, E.S., Langenheder, S., 2012. Local and regional factors influencing bacterial community assembly. Environ. Microbiol. Rep. 4, 1–9. https://doi.org/10.1111/j.1758-2229.2011.00257.x.

Liu, X., Nie, Y., Wu, X.L., 2023. Predicting microbial community compositions in wastewater treatment plants using artificial neural networks. Microbiome 11, 93. https://doi.org/10.1186/s40168-023-01519-9.

Matar, G.K., Bagchi, S., Zhang, K., Oerther, D.B., Saikaly, P.E., 2017. Membrane biofilm communities in full-scale membrane bioreactors are not randomly assembled and consist of a core microbiome. Water. Res. 123, 124–133. https://doi.org/10.1016/j.watres.2017.06.052.

McIlroy, S.J., Albertsen, M., Andresen, E.K., Saunders, A.M., Kristiansen, R., Stokholm-Bjerregaard, M., Nielsen, K.L., Nielsen, P.H., 2014. *Candidatus* Competibacter-lineage genomes retrieved from metagenomes reveal functional metabolic diversity. ISMe J. 8, 613–624. https://doi.org/10.1038/ismej.2013.162.

McIlroy, S.J., Onetto, C.A., McIlroy, B., Herbst, F.A., Dueholm, M.S., Kirkegaard, R.H., Fernando, E., Karst, S.M., Nierychlo, M., Kristensen, J.M., Eales, K.L., Grbin, P.R., Wimmer, R., Nielsen, P.H., 2018. Genomic and in situ analyses reveal the *Micropruina* spp. as abundant fermentative glycogen accumulating organisms in enhanced biological phosphorus removal systems. Front. Microbiol. 9, 1004. https://doi.org/10.3389/fmicb.2018.01004.

Nguyen Quoc, B., Armenta, M., Carter, J.A., Bucher, R., Sukapanpotharam, P., Bryson, S. J., Stahl, D.A., Stensel, H.D., Winkler, M.K.H., 2021a. An investigation into the optimal granular sludge size for simultaneous nitrogen and phosphate removal. Water. Res. 198, 117119. https://doi.org/10.1016/j.watres.2021.117119.

Nguyen Quoc, B., Wei, S., Armenta, M., Bucher, R., Sukapanpotharam, P., Stahl, D.A., Stensel, H.D., Winkler, M.K.H., 2021b. Aerobic granular sludge: impact of size distribution on nitrification capacity. Water. Res. 188, 116445. https://doi.org/10.1016/j.watres.2020.116445.

Nielsen, P.H., Mielczarek, A.T., Kragelund, C., Nielsen, J.L., Saunders, A.M., Kong, Y., Hansen, A.A., Vollertsen, J., 2010. A conceptual ecosystem model of microbial communities in enhanced biological phosphorus removal plants. Water Res. 44, 5070–5088. https://doi.org/10.1016/j.watres.2010.07.036.

Nielsen, P.H., Thomsen, T.R., Nielsen, J.L., 2004. Bacterial composition of activated sludge - importance for floc and sludge properties. Water Sci. Technol. 49, 51–58. https://doi.org/10.2166/wst.2004.0606.

Oksanen, J., Simpson, G.L., Guillaume, B., F., Kindt, R., Legendre, Pierre, Minchin, Peter R., O'Hara, R.B., Solymos, P., Stevens, Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De-Caceres, M., Durand, S., Evangelista, H., 2022. vegan: community ecology package.

Pang, Z., Lu, Y., Zhou, G., Hui, F., Xu, L., Viau, C., Spigelman, A.F., MacDonald, P.E., Wishart, D.S., Li, S., Xia, J., 2024. MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. Nucleic. Acids. Res. 52, W398–W406. https://doi.org/10.1093/nar/gkae253.

Pronk, M., De Kreuk, M.K., De Bruin, B., Kamminga, P., Kleerebezem, R., Van Loosdrecht, M.C.M., 2015. Full scale performance of the aerobic granular sludge process for sewage treatment. Water. Res. 84, 207–217. https://doi.org/10.1016/j.watres.2015.07.011.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic. Acids. Res. 41, D590–D596. https://doi.org/10.1093/nar/gks1219.

Richter-Heitmann, T., Hofner, B., Krah, F.S., Sikorski, J., Wüst, P.K., Bunk, B., Huang, S., Regan, K.M., Berner, D., Boeddinghaus, R.S., Marhan, S., Prati, D., Kandeler, E., Overmann, J., Friedrich, M.W., 2020. Stochastic dispersal rather than deterministic selection explains the spatio-temporal distribution of soil bacteria in a temperate grassland. Front. Microbiol. 11, 1391. https://doi.org/10.3389/fmicb.2020.01391.

Robeson, M.S., O'Rourke, D.R., Kaehler, B.D., Ziemski, M., Dillon, M.R., Foster, J.T., Bokulich, N.A., 2020. RESCRIPt: reproducible sequence taxonomy reference database management for the masses. https://doi.org/10.1101/2020.10.05.326504.

Ruiz-Haddad, L., Ali, M., Pronk, M., Van Loosdrecht, M.C.M., Saikaly, P.E., 2024. Demystifying polyphosphate-accumulating organisms relevant to wastewater treatment: a review of their phylogeny, metabolism, and detection. Environ. Sci. Ecotechnol. 100387. https://doi.org/10.1016/j.ese.2024.100387.

Saunders, A.M., Albertsen, M., Vollertsen, J., Nielsen, P.H., 2016. The activated sludge ecosystem contains a core community of abundant organisms. ISMe J. 10, 11–20. https://doi.org/10.1038/ismej.2015.117.

Singleton, C.M., Petriglieri, F., Wasmund, K., Nierychlo, M., Kondrotaite, Z., Petersen, J. F., Peces, M., Dueholm, M.S., Wagner, M., Nielsen, P.H., 2022. The novel genus, *Candidatus* Phosphoribacter, previously identified as *Tetrasphaera*, is the dominant polyphosphate accumulating lineage in EBPR wastewater treatment plants worldwide. ISMe J. 16, 1605–1616. https://doi.org/10.1038/s41396-022-01212-z.

Sravan, J.S., Matsakas, L., Sarkar, O., 2024. Advances in biological wastewater treatment processes: focus on low-carbon energy and resource recovery in biorefinery context. bioengineering 11, 281. https://doi.org/10.3390/bioengineering11030281.

Świątczak, P., Cydzik-Kwiatkowska, A., 2018. Performance and microbial characteristics of biomass in a full-scale aerobic granular sludge wastewater treatment plant. Environ. Sci. Pollut. Res. 25, 1655–1669. https://doi.org/10.1007/s11356-017-0615-9.

Van Dijk, E.J.H., Haaksman, V.A., Van Loosdrecht, M.C.M., Pronk, M., 2022. On the mechanisms for aerobic granulation - model based evaluation. Water. Res. 216, 118365. https://doi.org/10.1016/j.watres.2022.118365.

Vestergaard, S.Z., Peces, M., Murguz, A., Dueholm, M.K.D., Nierychlo, M., Nielsen, P.H., 2024. Microbial core communities in activated sludge plants are strongly affected by immigration and geography. Environ. Microbiome 19, 63. https://doi.org/10.1186/s40793-024-00604-2.

Weissbrodt, D.G., Neu, T.R., Kuhlicke, U., Rappaz, Y., Holliger, C., 2013. Assessment of bacterial and structural dynamics in aerobic granular biofilms. Front. Microbiol. 4. https://doi.org/10.3389/fmicb.2013.00175.

Winkler, M.K.H., Kleerebezem, R., Strous, M., Chandran, K., Van Loosdrecht, M.C.M., 2013. Factors influencing the density of aerobic granular sludge. Appl. Microbiol. Biotechnol. 97, 7459–7468. https://doi.org/10.1007/s00253-012-4459-4.

Winkler, M.K.H., Meunier, C., Henriet, O., Mahillon, J., Suárez-Ojeda, M.E., Del Moro, G., De Sanctis, M., Di Iaconi, C., Weissbrodt, D.G., 2018. An integrative review of granular sludge for the biological removal of nutrients and recalcitrant organic matter from wastewater. Chem. Eng. J. 336, 489–502. https://doi.org/10.1016/j.cej.2017.12.026.

Wu, Linwei, Ning, D., Zhang, B., Li, Y., Zhang, P., Shan, X., Zhang, Qiuting, Brown, M.R., Li, Z., Van Nostrand, J.D., Ling, F., Xiao, N., Zhang, Ya, Vierheilig, J., Wells, G.F., Yang, Y., Deng, Y., Tu, Q., Wang, A., , Global Water Microbiome Consortium, Acevedo, D., Agullo-Barcelo, M., Alvarez, P.J.J., Alvarez-Cohen, L., Andersen, G.L., De Araujo, J.C., Boehnke, K.F., Bond, P., Bott, C.B., Bovio, P., Brewster, R.K., Bux, F., Cabezas, A., Cabrol, L., Chen, S., Criddle, C.S., Deng, Y., Etchebehere, C., Ford, A., Frigon, D., Sanabria, J., Griffin, J.S., Gu, A.Z., Habagil, M., Hale, L., Hardeman, S.D., Harmon, M., Horn, H., Hu, Z., Jauffur, S., Johnson, D.R., Keller, J., Keucken, A., Kumari, S., Leal, C.D., Lebrun, L.A., Lee, J., Lee, M., Lee, Z.M.P., Li, Y., Li, Z., Li, M., Li, X., Ling, F., Liu, Y., Luthy, R.G., Mendonça-Hagler, L.C., De Menezes, F.G.R., Meyers, A.J., Mohebbi, A., Nielsen, P.H., Ning, D., Oehmen, A., Palmer, A., Parameswaran, P., Park, J., Patsch, D., Reginatto, V., De Los Reyes, F.L., Rittmann, B. E., Noyola, A., Rossetti, S., Shan, X., Sidhu, J., Sloan, W.T., Smith, K., De Sousa, O.V., Stahl, D.A., Stephens, K., Tian, R., Tiedje, J.M., Tooker, N.B., Tu, Q., Van Nostrand, J.D., De Los Cobos Vasconcelos, D., Vierheilig, J., Wagner, M., Wakelin, S., Wang, A., Wang, B., Weaver, J.E., Wells, G.F., West, S., Wilmes, P., Woo, S.G., Wu, Linwei, Wu, J.H., Wu, Liyou, Xi, C., Xiao, N., Xu, M., Yan, T., Yang, Y., Yang, M., Young, M., Yue, H., Zhang, B., Zhang, P., Zhang, Qiuting, Zhang, Ya, Zhang, T., Zhang, Qian, Zhang, W., Zhang, Yu, Zhou, H., Zhou, J., Wen, X., Curtis, T.P., He, Q., He, Z., Brown, M.R., Zhang, T., He, Z., Keller, J., Nielsen, P.H., Alvarez, P.J.J., Criddle, C.S., Wagner, M., Tiedje, J.M., He, Q., Curtis, T.P., Stahl, D.A., Alvarez-Cohen, L., Rittmann, B.E., Wen, X., Zhou, J., 2019. Global diversity and biogeography of bacterial communities in wastewater treatment plants. Nat. Microbiol. 4, 1183–1195. https://doi.org/10.1038/s41564-019-0426-5.

Xia, J., Ye, L., Ren, Hongqiang, Zhang, X.X., 2018. Microbial community structure and function in aerobic granular sludge. Appl. Microbiol. Biotechnol. 102, 3967–3979. https://doi.org/10.1007/s00253-018-8905-9.

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., Glöckner, F.O., 2014. The SILVA and "all-species living tree project (LTP)" taxonomic frameworks. Nucleic. Acids. Res. 42, D643–D648. https://doi.org/10.1093/nar/gkt1209.

Zhong, S., Hou, B., Zhang, J., Wang, Y., Xu, X., Li, B., Ni, J., 2023. Ecological differentiation and assembly processes of abundant and rare bacterial subcommunities in karst groundwater. Front. Microbiol. 14, 1111383. https://doi.org/10.3389/fmicb.2023.1111383.